

# On Dependence Structure of Copula-based Markov chains

Martial Longla

University of Cincinnati

Department of mathematical sciences

PO Box 210025, Cincinnati, Oh 45221-0025, USA.

martiala@mail.uc.edu

July 25, 2012

## Abstract

Dependence coefficients have been widely studied for Markov processes defined by a set of transition probabilities and an initial distribution. In this paper, we provide new tools to check the convergence rates of copula-based Markov chains. We also improve some known results of Beare and provide a necessary condition for Markov chains based on Archimedian copulas to be exponential  $\rho$ -mixing. Two general necessary condition on copulas to generate exponential rho mixing are given.

Key words: Markov chains, copula, mixing conditions, reversible processes, ergodicity.

AMS 2000 Subject Classification: Primary 60J20, 60J35, 37A30.

## 1 Introduction

This work is motivated by a paper by Xiaohong Chen, Wei Biao Wu and Yanping Yi and a result of B.K. Beare. In the paper [12], it was shown that Markov processes generated by the Clayton, Gumbel or Student copulas are geometrically ergodic. They used in their paper quantile transformations and small sets to show geometric ergodicity, but could not handle for instance the mixture of these copulas. In the paper [7], we have shown that these examples are actually exponential rho mixing and also answer the open question on geometric ergodicity of convex combinations of geometrically ergodic Markov chains. But we failed to relax the conditions of Beare's theorem to obtain exponential  $\rho$ -mixing rate without  $\phi$ -mixing rate. This task is one of the major concerns of this paper.

Quantifying the dependence among two or more random variables has been an enduring task for statisticians. A rich set of dependence measures has been proposed, including the well-known Pearson's correlation coefficients, Kendall's  $\tau$  and Spearman's  $\rho$  for bivariate random variables. While these measures are simple and can be easily computed, they are designed to capture only certain aspects of dependence. Indeed, it is rather unreasonable to expect a single scalar measure to have the capability to quantify all the dependence existing among the random variables. Copulas are full measures of dependence among components of random vectors. Unlike marginal and joint distributions, which are directly observable, a copula is a hidden dependence structure that couples

a joint distribution with its marginals. An early statistical application of copulas can be found in [3], where the dependence between two survival times in a multiple events study is modeled by the so-called Clayton copula

$$C(x, y) = (x^{-\alpha} + y^{-\alpha} - 1)^{-1/\alpha} \quad \alpha \geq 0.$$

The literature on copulas is growing fast; an excellent overview, guide to the literature and applications can be found in [6]. In later research into copulas, a driving force has been in financial risk management for modeling dependence among different assets in a portfolio. [8] can be regarded as one of the best books for an introduction to copulas. We also define a new list of copula with functional parameters, that can be used in various model.

## 1.1 Definitions

### 1.1.1 2-Copulas

A 2-copula is a function  $C : [0, 1] \times [0, 1] \rightarrow [0, 1] = I$  with the following properties:

1.  $C(0, y) = C(x, 0) = 0$  - meaning that  $C$  is grounded;
2.  $C(1, x) = C(x, 1) = x$  - meaning that each coordinate is uniform on  $I$ ;
3. for all  $[x_1, x_2; y_1, y_2] \subset I^2$ .

$$C(x_1, y_1) + C(x_2, y_2) - C(x_1, y_2) - C(x_2, y_1) \geq 0.$$

The definition doesn't mention probability, but in fact these conditions imply that  $C$  is the joint cumulative distribution function of two random variables with margins uniform on  $I$ .

It follows from the definition (see [8, chapter 1]) that the function  $C$  is non-decreasing in each of the parameters, has partial derivatives a.e with values between 0 and 1. The partial derivative of  $C$  are non-decreasing in the other parameter. A convex combination of 2-copulas is a 2-copula. An  $m$ -copula is defined for  $m > 2$  by induction:

1.  $C(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_m) = 0 \quad \forall i$ ;
2.  $C(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_m)$  - is a  $(m-1)$ -copula  $\forall i$ ;
3.  $\sum_V C(V) \sigma(V) \geq 0$  For all rectangle in  $I^m$ , where the sum is taken over all vertices of the rectangle, and  $\sigma(V) = (-1)^n$ ,  $n$ - number of initial components of the basic intervals of the rectangle present at  $V$ .

### 1.1.2 Sklar's Theorem

If  $X_1, \dots, X_m$  are random variables with joint distribution  $F$  and marginal distributions  $F_1, \dots, F_m$ , then the function

$$C(F_1(x_1), \dots, F_m(x_m)) = F(x_1, \dots, x_m)$$

is a copula with the properties given above. Moreover, if the random variables are continuous, then the copula is uniquely defined by the joint distribution and the marginal distributions by the formula

$$F(F_1^{-1}(x_1), \dots, F_m^{-1}(x_m)) = C(x_1, \dots, x_m).$$

The implication of the Sklar's Theorem is that, after standardizing the effects of marginals, the dependence among components of  $X$  is fully described by the copula. Indeed, most conventional measures of dependence can be explicitly expressed in terms of the copula.

*"This shows that the information concerning dependence properties is actually hidden in the copula. So, they present a natural framework for many investigations"- Nielsen.*

### 1.1.3 Copulas and Markov processes

Copulas have been shown to be a more flexible way to define a Markov process, as in the case when one suspects that the marginal distributions of the states are not related to the distribution of the initial state. Using copulas will allow changes in single marginal distributions, without having to change all other distributions in the chain.

A stationary Markov chain can be defined by a copula and a one dimensional marginal distribution. For stationary Markov chains, the transition probabilities are equal to the derivative of the copula with respect to the parameter we condition on (for more details, see [4, theorem 3.1]). This relationship was used in [12] to show that stationary Markov processes defined by the Clayton, Gumbel or Student copulas are geometrically ergodic. These notions will be defined later on.

We will use in this paper the following conventional notation:

### 1.1.4 Notations

1.  $L_f^2(\chi) = \{g(x) : \int_\chi g^2(x)f(x)dx < \infty\};$
2.  $L_0^2(\chi) = \{g(x) : \int_\chi g^2(x)f(x)dx < \infty, \int_\chi g(x)f(x)dx = 0\};$
3.  $\|g\|_2^2 = \int_I g^2(x)dx;$
4.  $A_{,i}(x_1, x_2) = \frac{\partial A(x_1, x_2)}{\partial x_i};$
5.  $c(x, y)$ - density of  $C(x, y)$ .

### 1.1.5 Dependence coefficients

Many dependence coefficients have been studied in the literature, such as  $\alpha_n$ ,  $\beta_n$ ,  $\rho_n$ ,  $\phi_n$  among others. In our paper, we will mainly use the last 3 coefficients, that are defined as follows:

Given  $\sigma$ -fields  $\mathbb{A}, \mathbb{B}$ :

$$\begin{aligned}\beta(\mathbb{A}, \mathbb{B}) &= \mathbb{E} \sup_{B \in \mathbb{B}} |P(B|\mathbb{A}) - P(B)| \\ \rho(\mathbb{A}, \mathbb{B}) &= \sup_{f \in L^2(\mathbb{A}), g \in L^2(\mathbb{B})} \text{corr}(f, g), \\ \phi_n(\mathbb{A}, \mathbb{B}) &= \sup_{B \in \mathbb{B}, A \in \mathbb{A}, P(A) > 0} |P(B|A) - P(B)|\end{aligned}$$

Given the alternative form of the transition probabilities for a Markov chain generated by a copula and a marginal distribution with positive density, it was shown in [7] that these coefficients have the following simple form when the copula is absolutely continuous with density  $c$ :

$$\begin{aligned}\beta_n &= \int_0^1 \sup_{B \subset I} \left| \int_B (c_n(x, y) - 1) dy \right| dx, \\ \phi_n &= \sup_{B \subset I} \text{ess sup}_x \left| \int_B (c_n(x, y) - 1) dy \right|, \\ \rho_n &= \sup \left\{ \int_0^1 \int_0^1 c_n(x, y) f(x) g(y) dx dy : \|g\|_2 = \|f\|_2 = 1, \mathbb{E}(f) = \mathbb{E}(g) = 0 \right\}.\end{aligned}$$

Here,  $c_n$  is the density of the random variable  $(X_0, X_n)$ . In general the following inequalities hold (see [2, theorem 7.4, 7.5] for more.):

$$\beta_n \leq \phi_n, \quad \rho_n \leq 2\sqrt{\phi_n}, \quad \rho_n \leq (\rho_1)^n. \quad (1)$$

These coefficients are defined to assess the dependence structure of the Markov process, and provide necessary conditions for CLT and functional CLT necessary for the statistics based on the studied model. Some examples can be found in [9], [10] and the references therein. A stochastic process is said to be respectively an  $\alpha$ -mixing ( $\beta$ -mixing or  $\rho$ -mixing), if  $\alpha_n \rightarrow 0$  ( $\beta_n \rightarrow 0$  or  $\rho_n \rightarrow 0$ ). The process is exponentially mixing, if the convergence rate is exponential. A stochastic process is said to be geometrically ergodic, if for some  $a \in (0, 1)$ , and  $n \in \mathbb{N}$ ,

$$\sup_{B \subset \mathcal{X}} |P(X_n \in B | X_0 = x) - P(X_n \in B)| \leq a^n W(x), \quad \text{with } \mathbb{E}(W(X_0)) < \infty.$$

By theorem 2.1 of Nemmelin and Tuominen (1982), *geometric ergodicity is equivalent to exponential  $\beta$ -mixing*. So, whenever we encounter the term, we may understand either of the two concepts. The main results on this topic are due to Darsow, Nguyen and Olsen (1992), de la Peña, Ibragimov and Sharakhmetov (2006) and Ibragimov (2009), who were among the first to study copula-based Markov processes. They characterized copula-based Markov chains. Joe (1997) proposed a class of parametric (strictly) stationary Markov models based on parametric copulas and parametric invariant distributions. Their setup of the problem was modified by Chen and Fan (2006), who studied a class of semi-parametric stationary Markov models based on parametric copulas and nonparametric invariant distributions and analyzed the strength of dependence in the Markov chain. They showed that the temporal dependence measure is fully determined by the properties of the copulas. Studying the strength of the dependence, Beare (2008) provided simple sufficient conditions for geometric  $\beta$ -mixing in terms of copulas without any tail dependence. None of these papers is able to verify whether or not a Markov process generated via a tail dependent copula such as Clayton or others, is geometric  $\rho$ -mixing.

The tail dependent copulas are the most popular copulas because they suitably model shocks in economics. In an effort to address this issue, Ibragimov and Lentzas (2008) demonstrated via simulation that Clayton copula-based first-order strictly stationary Markov models could behave as “long memory” in copula levels (meaning that dependence on the initial distribution tends to persist). But, Xiaohong Chen, Wei Biao Wu and Yanqing Yi (2009) showed that this simulation is misleading, because the copula-based models studied by Ibragimov are actually geometric  $\beta$ -mixing. In conclusion, although a time series plot may look highly persistent and “long memory alike,” the time series may still be weakly dependent and “short memory ” so simulating isn’t a good way to handle the issue.

Recently in a discussion with Beare, I learned he had published a result improving the results in [12]. Beare provided in a recent paper the proof based on simulation, that the Clayton copula provides exponential  $\rho$ -mixing. The theoretical proof of this fact and many others was given by M. Longla and M. Peligrad in [7].

The paper is structured as follows: In section two below we provide new results on exponential  $\rho$ -mixing, and exponential  $\beta$ -mixing for some families of copulas that were not studied before and some new families of copulas that generate exponential  $\rho$ -mixing and  $\phi$ -mixing are given. In section 3 we provide the proofs of all theorems.

## 2 Results on dependence coefficients

We shall recall that exponential  $\rho$ -mixing is equivalent to  $\rho_1 < 1$  for Markov chains. This fact will be used in this section to improve some known results.

## 2.1 General condition for exponential $\rho$ -mixing

**Theorem 1** *A copula-based Markov chain generated by a symmetric copula with square integrable density is geometric  $\beta$ -mixing if it is geometric  $\rho$ -mixing.*

This means, that for a Markov chain generated by a copula with square integrable density and an increasing marginal distribution, exponential  $\rho$ -mixing implies geometric ergodicity.

**Theorem 2** *Let  $c$  be the density of the copula  $C$ . Assume*

$$\lim_{y \rightarrow 0} yc(x, y) = \lim_{y \rightarrow 1} (1 - y)c(x, y) = 0; \quad (2)$$

$$\int_0^1 |c_y(x, y)| dy \in L_2([0, 1]), \quad \left\| \int_0^1 |c_y(x, y)| dy \right\|_2^2 = k_2; \quad (3)$$

$$c(x, 1) - c(x, 0) \in L_2([0, 1]), \quad \text{and} \quad \|c(x, 1) - c(x, 0)\| + \int_0^1 |c_y(x, y)| dy \Big|_2^2 = k_1 < \infty. \quad (4)$$

*If  $k_1 + k_2 < 12$ , then the Markov process generated by  $C$  is an exponential  $\rho$ -mixing ( $\rho_1 \leq \sqrt{(k_1 + k_2)/12} < 1$ ), and therefore geometrically ergodic if the chain is reversible.*

### 2.1.1 Example of the Farlie-Gumbel-Morgenstern family of copulas

$C(x, y) = xy + \theta xy(1 - x)(1 - y)$ ,  $c(x, y) = 1 + \theta(1 - 2x)(1 - 2y)$ ,  $c_y(x, y) = -2\theta(1 - 2x)$ ,  $\theta \in [-1, 1]$ . For this family, all assumptions of the theorem are met.

$$c(x, 1) - c(x, 0) = -2\theta(1 - 2x), \quad \int_0^1 |c_y(x, y)| dy = 2|\theta(1 - 2x)|;$$

$$\|c(x, 1) - c(x, 0)\| + \int_0^1 |c_y(x, y)| dy = 4|\theta(1 - 2x)|.$$

Therefore,  $k_1 = 4\theta^2/3$  and  $k_2 = 16\theta^2/3$ .  $k_1 + k_2 < 12$  if  $\theta^2 < 9/5$ . This is true for all  $\theta \in [-1, 1]$ . Therefore, the Farlie-Gumbel-Morgenstern family generates exponential  $\rho$ -mixing and an exponential  $\beta$ -mixing for all its parameters.

Beare proved that for a copula with density bounded away from zero we have exponential  $\rho$ -mixing. These conditions imply actually a stronger mixing condition ( $\phi$ ) as shown in [7]. This conditions are relaxed in the following theorem:

**Theorem 3** *If the copula of the Markov process is such that there exists nonnegative functions  $\varepsilon_1, \varepsilon_2$  defined on  $[0, 1]$  for which the density of the absolute continuous part of the copula denoted  $c(x, y)$  satisfies the inequality*

$$c(x, y) \geq \varepsilon_1(x) + \varepsilon_2(y)$$

*with  $\varepsilon_1, \varepsilon_2 \in L_1[0, 1]$  such that at least one of the two functions has a non-zero integral, then the process is exponential  $\rho$ -mixing, and thus exponential  $\beta$ -mixing if the density appears to be positive on a set of measure 1 or square integrable.*

**Remark 4** *The presence of the two functions in the theorem is just to emphasize that it doesn't matter which of them ends up providing a lower bound on the absolute continuous part of the density. The question is irrelevant if the copula is symmetric, but is an issue in general. To apply the theorem, one just needs to find any positive function of one variable that can allow a lower bound on the density with the stated property.*

**Remark 5** *This theorem improves Beare's result [1, Theorem 4.2], by extending it to cases when the density can actually be equal to zero on a set of non-zero measure, and therefore not be bounded away from 0. This is an improvement that avoids the conditions of Beare that allowed to show  $\phi$ -mixing rate in [7]. All examples from his paper are therefore covered by this theorem, and in the example below, we can exhibit a copula that provides exponential  $\rho$ -mixing, but was ruled out by his assumptions.*

## 2.2 Examples

1. The Plankett copula family is defined by

$$C(u, v) = \frac{1 + (\theta - 1)(u + v) - \sqrt{(1 + (\theta - 1)(u + v))^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)} \quad (5)$$

where  $\theta$  is the constant cross product ratio

$$\theta = \frac{P(U > u, V > v)P(U \leq u, V \leq v)}{P(U > u, V \leq v)P(U \leq u, V > v)}$$

and the density is given by

$$\begin{aligned} c(u, v) = & \frac{1}{2} \frac{(1 + \theta) \left( (1 + (\theta - 1)(u + v))^2 - 4uv\theta(\theta - 1) \right)}{\left( \sqrt{(1 + (\theta - 1)(u + v))^2 - 4uv\theta(\theta - 1)} \right)^3} + \\ & + \frac{1}{2} \frac{\left( (\theta - 1)(u + v) + 1 - 2u\theta \right) \left( (\theta - 1)(u + v) + 1 - 2v\theta \right)}{\left( \sqrt{(1 + (\theta - 1)(u + v))^2 - 4uv\theta(\theta - 1)} \right)^3}. \end{aligned} \quad (6)$$

For  $\theta \geq 1$ , The denominator of this density is less than or equal to  $1 + (\theta - 1)(u + v) \leq 1 + (\theta - 1)(1 + u)$ . The numerator is a polynomial in  $v$  with non-zero minimum in  $v$  for fixed values of  $u$ . Therefore, this density is bounded from below by an integrable function of  $u$ .

For  $\theta < 1$ , we can divide into two cases:  $\theta \leq 1/2$  and  $1/2 < \theta < 1$ . In both cases, it is easy to bound the denominator from above and the numerator from below to get a lower bound on the density of the copula. Thus, by theorem 3, this copula family provides an exponential  $\rho$ -mixing for all values of the parameter.

On this example, one can actually see the power of this theorem. For  $\theta > 1$  the density is not bounded from above (it would therefore be  $\phi$ -mixing), it is not square integrable (it would therefore fall in the range of theorem 1) and it is not bounded away from 0 (it would therefore fall in the range of Beare's theorem).

2. Given any functions  $h : [0, 1] \rightarrow [0, 1]$ ,  $g : [0, 1] \rightarrow [0, 1]$ , the function

$$c(x, y) = \frac{\sup g - g(x)h(y) + h(y)\|g\|_1 + g(x)\|h\|_1}{\sup g + \|g\|_1\|h\|_1}$$

is the density of a copula that generates an exponential  $\rho$ -mixing Markov chain.

(The proof uses theorem 3 with  $\varepsilon_1(x) = \frac{g(x)\|h\|_1}{\sup g + \|g\|_1\|h\|_1}$ ,  $\varepsilon_2(y) = \frac{h(y)\|g\|_1}{\sup g + \|g\|_1\|h\|_1}$ ). We have

$$\rho_1 \leq 1 - \frac{1}{2} \frac{2\|g\|_1\|h\|_1}{\sup g + \|g\|_1\|h\|_1} = \frac{\sup g}{\sup g + \|g\|_1\|h\|_1} < 1 \quad \text{using Theorem 3.}$$

3. Given any functions  $h : [0, 1] \rightarrow [0, 1]$ ,  $g : [0, 1] \rightarrow [0, 1]$ , the function

$$c(x, y) = \frac{\sup g \sup h - g(x)h(y) + h(y)\|g\|_1 + g(x)\|h\|_1}{\sup g \sup h + \|g\|_1\|h\|_1}$$

is the density of a copula that also defines an exponential  $\rho$ -mixing Markov chain. We have

$$\rho_1 \leq 1 - \frac{1}{2} \frac{2\|g\|_1\|h\|_1}{\sup g \sup h + \|g\|_1\|h\|_1} = \frac{\sup g \sup h}{\sup g \sup h + \|g\|_1\|h\|_1} < 1.$$

4. Given any functions  $h : [0, 1] \rightarrow [0, 1]$ ,  $g : [0, 1] \rightarrow [0, 1]$ , the function

$$c(x, y) = \frac{\sup g(\sup h - \inf h) - g(x)(\sup h - h(y)) + (\sup h - h(y))\|g\|_1 + g(x)(\sup h - \|h\|_1)}{\sup g(\sup h - \inf h) + \|g\|_1(\sup h - \|h\|_1)}$$

is the density of a copula that also defines an exponential  $\rho$ -mixing Markov chain. We have

$$\rho_1 \leq \frac{\sup g(\sup h - \inf h)}{\sup g(\sup h - \inf h) + \|g\|_1(\sup h - \|h\|_1)} < 1.$$

5. Given any functions  $h : [0, 1] \rightarrow [0, 1]$ ,  $g : [0, 1] \rightarrow [0, 1]$ , the function

$$c(x, y) = \frac{(\sup g - \inf g)(\sup h - \inf h) - (\sup g - g(x))(\sup h - h(y))}{(\sup g - \inf g)(\sup h - \inf h) + (\sup g - \|g\|_1)(\sup h - \|h\|_1)} +$$

$$+ \frac{(\sup h - h(y))(\sup g - \|g\|_1) + (\sup g - g(x))(\sup h - \|h\|_1)}{(\sup g - \inf g)(\sup h - \inf h) + (\sup g - \|g\|_1)(\sup h - \|h\|_1)}$$

is the density of a copula that also defines an exponential  $\rho$ -mixing Markov chain. We have

$$\rho_1 \leq \frac{(\sup g - \inf g)(\sup h - \inf h)}{(\sup g - \inf g)(\sup h - \inf h) + (\sup g - \|g\|_1)(\sup h - \|h\|_1)} < 1.$$

6. The following density also generates an exponential  $\rho$ -mixing with  $\rho_1 \leq \frac{1 + \frac{1}{2^{1-\alpha}}}{1 + \frac{3}{2^{1-\alpha}}} < 1$ .

$$c(x, y) = \frac{\frac{3}{2^{2-\alpha}} + 1 + (1/2 - y)x^{1/\alpha-1}\text{sign}(1/2 - x^{1/\alpha})}{1 + \frac{3}{2^{2-\alpha}}}, \quad \alpha \in (0, 1] \quad (7)$$

This density is bounded away from 0, thus it is also providing  $\phi$ -mixing.

7. Finally, the bivariate density function

$$c(x, y) = 1 + \frac{\theta}{2\alpha} x^{1/\alpha-1} (2y - 1) \text{sign}(1/2 - x^{1/\alpha}) \quad \theta \in [-2\alpha, 2\alpha], \quad \alpha \in (0, 1]$$

generates exponential  $\phi$ -mixing for  $|\theta| < 2\alpha$ .

In the last inequality of the proof of theorem 3, we obtain a guarantee of the following result:

**Lemma 6** *Let  $f(x, y)$  be a nonnegative function in  $L_1[0, 1]$  with the following properties:*

1.  $\int_I f(x, y) dx = 1$  a.e;
2.  $\int_I f(x, y) dy = 1$  a.e;

*Then, if there exist nonnegative functions  $\varepsilon_1$  and  $\varepsilon_2$  in  $L_1[0, 1]$  such that  $f(x, y) \geq \varepsilon_1(x) + \varepsilon_2(y)$  a.e, then*

$$\int_I \varepsilon_1(x) dx + \int_I \varepsilon_2(x) dx < 2 \quad \text{and}$$

$$\int_{I^2} f(x, y) g(x) h(y) dx dy \leq \left(1 - \frac{1}{2} \left( \int_I \varepsilon_1(x) dx + \int_I \varepsilon_2(x) dx \right) \right) \left( \int_I g^2(x) dx \right)^{1/2} \left( \int_I h^2(x) dx \right)^{1/2}.$$

## 2.3 Exponential $\rho$ -mixing for Archimedean copulas

### 2.3.1 Archimedean copulas

Archimedean copulas have been studied by many researchers and are very popular because of the associativity property that they share. A result by Beare (2010) is the first about general necessary conditions for exponential  $\rho$ -mixing. He proved that under some mild conditions, some strict Archimedean copulas generate exponential  $\rho$ -mixing.

Archimedean copulas are defined in [8] as follows: Given a function  $\varphi : [0, 1] \rightarrow [0, \infty)$  such that:  $\varphi$  is a decreasing concave up function (positive second derivative whenever it exists),  $\varphi(1) = 0$ ; if  $\varphi(0) = \infty$ , then  $\varphi$  is called a strict generator or generator of a strict Archimedean copula given by the formula:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)); \quad c(u, v) = -\frac{\varphi'' \circ \varphi^{-1}(\varphi(u) + \varphi(v)) \varphi'(u) \varphi'(v)}{\left(\varphi' \circ \varphi^{-1}(\varphi(u) + \varphi(v))\right)^3}$$

If  $\varphi(0) < \infty$ , then  $\varphi$  is called non-strict generator or generator of a non-strict Archimedean copula defined by

$$C(u, v) = \varphi^{-1}(\min(\varphi(u) + \varphi(v), \varphi(0))).$$

A non strict generator can always be standardized to have  $\varphi(0) = 1$ . It can be proved that all generators of the same Archimedean copula are scalar multiples of the standard generator (I will provide a proof of this in the appendix). So, without loss of generality I can state all my results with the standard generator, and convert later on to any initial generator.

**Theorem 7** *Let  $\varphi$  be a non-strict standard generator of an Archimedean copula not equal to the Hoeffding lower bound. Assume  $\varphi$  is twice differentiable and its second derivative is decreasing. Then, the copula generates an exponential  $\rho$ -mixing Markov process if one of the following inequalities is satisfied:*

$$\int_0^1 (1-x) \left( \frac{h(x)}{(\varphi' \circ \varphi^{-1}(x))^2} \right)^2 dx < 1, \quad \text{or} \quad \int_0^1 h^2(x)(1-x) dx < (\varphi'(1))^4 \quad \text{for } \varphi'(1) \neq 0$$

$$\text{where } h(x) = \max_{0 \leq y \leq 1-x} \varphi'' \circ \varphi^{-1}(x+y);$$

$h(x) = \varphi''(0)$  if  $\varphi''$  is decreasing and  $h(x) = \varphi'' \circ \varphi^{-1}(x)$  if  $\varphi''$  is increasing.

### 2.3.2 Example

The Archimedean copula with generator  $\tilde{\varphi}(u) = -\ln(\theta u + 1 - \theta)$ . The standard generator is

$$\varphi(x) = \frac{\ln(\theta u + 1 - \theta)}{\ln(1 - \theta)}, \quad \theta \in (0, 1).$$

$$\varphi^{-1}(x) = \frac{(1 - \theta)^x - 1 + \theta}{\theta}, \quad \varphi'(x) = \frac{1}{\ln(1 - \theta)(x + \frac{1-\theta}{\theta})}, \quad \varphi' \circ \varphi^{-1}(x) = \frac{\theta}{\ln(1 - \theta)(1 - \theta)^x},$$

$$\varphi'(1) = \frac{\theta}{\ln(1 - \theta)}, \quad \varphi'' \circ \varphi^{-1}(x+y) = \frac{\theta^2(1 - \theta)^{-2(x+y)}}{-\ln(1 - \theta)}, \quad h(x) = \frac{\theta^2(1 - \theta)^{-2}}{-\ln(1 - \theta)}.$$

$$\int_0^1 h^2(x)(1-x) dx = 1/2 \left( \frac{\theta^2(1 - \theta)^{-2}}{\ln(1 - \theta)} \right)^2 < (\varphi'(1))^4 = \left( \frac{\theta}{\ln(1 - \theta)} \right)^4;$$



$$\int_0^1 (1-x) \left( \frac{h(x)}{(\varphi' \circ \varphi^{-1}(x))^2} \right)^2 dx = -\frac{\ln(1-\theta)}{4(1-\theta)^4} + \frac{1}{16} - \frac{1}{16(1-\theta)^4} < 1.$$

So, we obtain exponential  $\rho$ -mixing for  $\theta \in (0, \theta_0)$ , where  $\theta_0 \cong .348$  is the unique value of  $\theta$  for which the inequality becomes an equality.

### 2.3.3 Remark

This example is taken from [8, §6], from the list of Archimedean copulas that have been used in applications. On this example we can see that the theorem doesn't handle the case of the independence copula, corresponding to  $\theta = 1$ . This is due to the fact that in this limiting case, the copula becomes strict. We might be tempted to include the case  $\theta = 0$  by only looking at the final inequality. A closer look at the original generator shows that for this value, the function is constant! On the other hand, the limiting copula is the Hoeffding lower bound, which is ruled out by the assumptions of the theorem.

### 2.3.4 Example

$$\varphi(x) = \frac{1-x}{1+(\theta-1)x} = \varphi^{-1}(x), \quad \theta \in [1, \infty)$$

$$\varphi'(x) = \frac{-\theta}{(1+(\theta-1)x)^2}, \quad \varphi' \circ \varphi^{-1}(x) = \frac{(1+(\theta-1)x)^2}{-\theta}, \quad \varphi''(x) = \frac{2\theta(\theta-1)}{(1+(\theta-1)x)^3}.$$

Again, we can see that the case  $\theta = 1$  is ruled out, as  $\varphi''(x) \equiv 0$  for this value.  $\varphi$  is decreasing, so,

$$h(x) = 2\theta(\theta-1), \quad \int_0^1 \left( \frac{(1-x)^{1/2} h(x)}{(\varphi' \circ \varphi^{-1}(x))^2} \right)^2 dx = 4\theta^6(\theta-1)^2 \int_0^1 \frac{(1-x)dx}{(1+(\theta-1)x)^8} = 4\theta^6 \int_1^\theta \frac{\theta-t}{t^8} dt;$$

$$\rho^2 \leq \frac{2}{21} + \frac{4}{7}\theta^7 - \frac{2}{3}\theta^6 = f(\theta);$$

$f(\theta)$  is an increasing function on  $[1, \infty)$  with  $f(1) = 0$ ,  $f(\infty) = \infty$ . Thus, there exists a unique  $\theta_0 \cong 1.388$  for which  $f(\theta_0) = 1$ . Therefore, the copula generates exponential  $\rho$ -mixing Markov chain for  $\theta \in (1, \theta_0)$ .

## 2.4 Other results

In the paper [7], a result reads as follows:

**Lemma 8** *Any convex combination of geometrically ergodic reversible Markov chains is geometrically ergodic.*

Two simple applications of this lemma lead to the following corollaries:

**Corollary 9** *Given two Markov processes with transition kernels  $P$  and  $Q$ , If  $P$  is geometrically ergodic and  $Q$  is irreducible, reversible and has a spectral gap as operator in  $L^2(\chi, \mu)$ , then  $\alpha P + (1-\alpha)Q$  is geometrically ergodic and has a spectral gap in  $L^2(\chi, \mu)$ .*

**Corollary 10** *The Markov chain generated by the convex combination of symmetric copulas, whose Markov processes are geometrically ergodic is also geometrically ergodic and exponential  $\rho$ -mixing.*

### 2.4.1 Exponential $\beta$ -mixing for the Frechet and Mardia families of copulas

$$C(x, y) = \frac{\theta^2(1 + \theta)}{2}M(x, y) + (1 - \theta^2)P(x, y) + \frac{\theta^2(1 - \theta)}{2}W(x, y), \quad \theta \in [-1, 1] \quad (8)$$

$$C(x, y) = \alpha M(x, y) + (1 - \alpha - \beta)P(x, y) + \beta W(x, y) \quad (0 \leq \alpha + \beta \leq 1) \quad (9)$$

(8) defines the Mardia family of copulas and (9) defines the Frechet family of copulas. Notice that a Mardia copula is a Frechet copula with  $\alpha + \beta = \theta^2$ . The copulas  $W$  and  $M$  are the Hoeffding lower and upper bounds respectively and  $P$  is the independence copula. For any copula from this family, there exists a singular density

$$C_{,1}(u, v) = \alpha \mathbb{I}(u - v \leq 0) + \beta \mathbb{I}(1 - u - v \leq 0) + (1 - \alpha - \beta)v;$$

$$c(u, v) = \alpha \delta_u(v) + \beta \delta_{1-u}(v) + 1 - \alpha - \beta.$$

If the Markov chain is generated by a copula from the Frechet (Mardia) family with  $(\alpha + \beta \neq 1)$  ( $\theta \neq 1$ ), then the  $n$  steps joint distribution is given by

$$C^n(x, y) = C_{\alpha_n, \beta_n}(x, y); \quad \alpha_n = \frac{1}{2}[(\alpha + \beta)^n + (\alpha - \beta)^n], \beta_n = \frac{1}{2}[(\alpha + \beta)^n - (\alpha - \beta)^n] \quad (10)$$

and the process is exponentially  $\beta$ -mixing. For the excluded parameters there is no mixing.

**Remark 11** 1. First of all, notice that theorem 3 could be applied to both families to show that we have exponential  $\rho$ -mixing for cases when we have geometric ergodicity above; but the theorem wouldn't answer the question for the rest of the values of the parameters.

2. On this example as in the previous examples, we find the copula  $W$  as a limiting copula in the family, and there is exponential  $\rho$ -mixing for admissible parameters in the neighborhood of the parameter that corresponds to it. This opens some question for further research: What is the reason of such behavior? Is it always the case?

### 2.4.2 Practical Example for the Metropolis Hastings Algorithm

The popular kernel of the independent Metropolis Hastings Algorithm, that is used to generate Markov chains with a given probability of staying at the same state  $x$  equal to  $p(x)$ , where  $x \in [-1, 1]$ , is defined by:

$$Q(x, A) = p(x)\delta_A(x) + (1 - p(x))\mu(A), \quad \text{where } \mu \text{ is a measure on } [-1, 1].$$

If  $\theta = \int_{-1}^1 \frac{\mu(dx)}{1-p(x)} < \infty$ , then the invariant distribution of this process is  $\pi$  defined by

$$\pi(dx) = \frac{\mu(dx)}{\theta(1 - p(x))}.$$

Therefore, if we allow the acceptance probability to depend on a parameter  $\beta$ , and require the marginal distribution to be uniform on  $[-1, 1]$  and  $\mu$  absolutely continuous with respect to the Lebesgue measure with density  $h(x, \beta)$ , then we obtain the following:

$$Q(x, A) = p(x, \beta)\delta_A(x) + (1 - p(x, \beta))\mu(A);$$

$$\pi(dx) = \frac{1}{2}dx \Rightarrow h(x, \beta) = k(1 - p(x, \beta)), \quad \theta = 2k.$$

The corresponding copula representation is therefore defined by

$$C_{,1}\left(\frac{x+1}{2}, \frac{y+1}{2}\right) = p(x, \beta)\mathbb{I}(x \leq y) + k(1 - p(x, \beta)) \int_{-1}^y (1 - p(t, \beta))dt.$$

So, the derivative of the copula is

$$C_{,1}(u, v) = p(2u - 1, \beta)\mathbb{I}(u \leq v) + k(1 - p(2u - 1, \beta)) \int_{-1}^{2v-1} (1 - p(t, \beta))dt, \quad f(x) = \int_{-1}^x p(x, \beta)dt.$$

$$\text{Therefore, } C(u, v) = \frac{1}{2} \left[ f\left(\min(2u - 1, 2v - 1)\right) + k(2u - f(2u - 1))(2v - f(2v - 1)) \right], \quad (11)$$

For this process, if we take  $p(x, \beta) = \beta|x|$  with  $\beta < 1$ , then due to [5, Lemma 2],

$$\beta_n \leq 3\mathbb{E}_\pi(p^{n/2}(X, \beta)) = \frac{3\beta^{n/2+1}}{n/2+1} \leq k\rho^n, \quad \rho = \sqrt{\beta};$$

On the other hand,  $Q(x, \cdot)g(x) = p(x)g(x) = \beta|x|g(x)$  for  $g \in L_{2,0}(-1, 1)$  implies

$$\|Q\|_2 \leq \beta \left( \sup_{\|g\|_2=1} \int_{-1}^1 |x|^2 g^2(x) dx \right)^{1/2} \leq \beta \left( \sup_{\|g\|_2=1} \int_{-1}^1 g^2(x) dx \right)^{1/2} = \beta < 1.$$

Therefore, this process is both exponential  $\rho$ -mixing and exponential  $\beta$ -mixing. For  $\beta = 1$ , this process is a  $\beta$ -mixing with  $1/n$  decay rate, but fails to be  $\rho$ -mixing. In general, any function of the form (11) with an increasing differentiable function  $f$ , such that  $f(-1) = 0$ ,  $f(1) = 2 - 1/k$  will define a one parameter copula family with  $1 \geq k \geq 1/2$ . This family will provide an exponential  $\rho$ -mixing for  $1/2 \leq k < 1$ .

**Remark 12** *The Farlie-Gumbel Morgenstern's example provides actually exponential  $\phi$ -mixing when parameters are in the interior of the domain, and this theorem allows to prove the mixing rate for the boundary values. Theorem 3 could also be applied to this problem to equivalently obtain the same result for  $\theta \in (-1, 1)$ . Theorem 3 would not allow to investigate the boundary cases. This shows that these theorems are complementary, and do not follow from each other.*

## 3 Appendix 1: Mathematical proofs

### 3.1 Proof of theorem 1

The square integrable density defines a Hilbert-Schmidt operator, and therefore a compact operator. So, there exist a basis of eigen-functions in  $L_2(0, 1)$ . Reversibility implies a spectral representation of the operator in the form

$$c(u, v) = \sum_{i=0}^{\infty} \lambda_i \phi_i(u) \phi_i(v),$$

where  $\phi_i(u)$  are the eigen-functions corresponding to the eigen-values  $\lambda_i$ , and form an orthonormal basis of  $L_2$ . All  $\lambda_i$  exist and are positive real numbers in the decreasing order. 1 is eigen-value with eigen-function 1. Therefore,  $\lambda_i^k$  are eigen-values of the operator  $c^k$  with the same corresponding eigen-functions. So,

$$\begin{aligned}
c_k(u, v) &= 1 + \sum_{i=1}^{\infty} \lambda_i^k \phi_i(u) \phi_i(v), \\
\rho_k &= \sup \left\{ \left| \int_0^1 c_k(u, v) f(u) g(v) du dv \right| : \mathbb{E}(f) = \mathbb{E}(g) = 0, \mathbb{E}(f^2) = \mathbb{E}(g^2) = 1 \right\} \\
&= \left| \int_0^1 c_k(u, v) f(u) g(v) du dv \right| = \left| \sum_{i=1}^{\infty} \lambda_i^k \int_0^1 \phi_i(u) f(u) du \int_0^1 \phi_i(v) g(v) dv \right| \leq \\
&\leq \lambda_i^k \left( \int_0^1 \phi_i^2(u) du \right)^{1/2} \left( \int_0^1 f^2(u) du \right)^{1/2} \left( \int_0^1 \phi_i^2(v) dv \right)^{1/2} \left( \int_0^1 g^2(v) dv \right)^{1/2} \leq \sum_{i=1}^{\infty} \lambda_i^k.
\end{aligned}$$

Then, for  $k > 2$ , we have

$$\rho_k \leq \sum_{i=1}^{\infty} \lambda_i^k \leq \lambda_1^{k-2} \sum_{i=1}^{\infty} \lambda_i^2 \leq c \lambda_1^k.$$

Here  $c = (\sum_{i=1}^{\infty} \lambda_i^2) / \lambda_1^2$ ; and the series converges because we have a Hilbert-Schmidt operator.

Therefore, if  $\lambda_1 < 1$ , then  $\rho_k$  converges to 0 exponentially. On the other hand, if  $\lambda_1 = 1$ , then

$$c_k(u, v) = 1 + \lambda_1^k \phi_1(u) \phi_1(v) + \sum_{i=2}^{\infty} \lambda_i^k \phi_i(u) \phi_i(v);$$

Because the basis is orthonormal, we have  $\langle \phi_i(u), \phi_1(u) \rangle = \int_0^1 \phi_i(u) \phi_1(u) du = 0$  for  $i \neq 1$ ; and

$$\langle c_k(u, v) \phi_1(u), \phi_1(v) \rangle = \lambda_1^k \int_0^1 \phi_1^2(u) du \langle \phi_1(v), \phi_1(v) \rangle = \lambda_1^k.$$

Thus,  $1 \geq \rho_k \geq \lambda_1^k$ . Therefore, if  $\lambda_1 = 1$ , then  $\rho_k = 1$  for all  $k$ . So, we have exponential  $\rho$ -mixing if and only if  $\rho_1 < 1$ .

$$\text{Now, recall } \beta_k = \int_0^1 \sup_B \left| \int_B (c_k(u, v) - 1) du dv \right|. \text{ So,}$$

$$\begin{aligned}
\beta_k &= \int_0^1 \sup_B \left| \int_B \left( \sum_{i=1}^{\infty} \lambda_i^k \phi_i(u) \phi_i(v) \right) du dv \right| \leq \sum_{i=1}^{\infty} \lambda_i^k \int_0^1 |\phi_i(v)| dv \sup_B \left| \int_B \phi_i(u) du \right| \leq \\
&\leq \sum_{i=1}^{\infty} \lambda_i^k \int_0^1 |\phi_i(v)| dv \sup_B \mu^{1/2}(B) \left( \int_B \phi_i^2(u) du \right)^{1/2} \\
&\leq \sum_{i=1}^{\infty} \lambda_i^k \int_0^1 |\phi_i(v)| dv \leq \sum_{i=1}^{\infty} \lambda_i^k \left( \int_0^1 |\phi_i^2(v)| dv \right)^{1/2} = \sum_{i=1}^{\infty} \lambda_i^k.
\end{aligned}$$

Therefore,  $\beta_k \leq c \lambda_1^k$ . So,  $\beta_k$  converges exponentially to 0 when  $\rho_1 < 1$ .

### 3.2 Proof of theorem 3

Let  $f, g$  be two functions with  $\|f\|_2 = \|g\|_2 = 1, \mathbb{E}(f(X)) = \mathbb{E}(g(X)) = 0$ . We have  $2f(x)g(y) = f^2(x) + g^2(y) - (f(x) - g(y))^2$ . Therefore,

$$2 \int_{I^2} f(x)g(y)C(dx, dy) = \int_{I^2} f(x)^2 C(dx, dy) + \int_{I^2} g(y)^2 C(dx, dy) - \int_{I^2} (f(x) - g(y))^2 C(dx, dy).$$

$$\text{But, } \int_{I^2} f(x)^2 C(dx, dy) = \int_I f(x)^2 \int_I C(dx, dy) = \int_{I^2} f(x)^2 dy = 1 = \int_{I^2} g(y)^2 C(dx, dy),$$

$$\int_{I^2} (f(x) - g(y))^2 C(dx, dy) \geq \int_{I^2} (f(x) - g(y))^2 (\varepsilon_1(x) + \varepsilon_2(y)) dx dy =$$

$$= \int_{I^2} (f^2(x) + g^2(y) - 2f(x)g(y))(\varepsilon_1(x) + \varepsilon_2(y)) dx dy = I_a + I_b$$

$$I_b = \int_{I^2} f^2(x)\varepsilon_2(y) dx dy + \int_{I^2} g^2(y)\varepsilon_2(y) dx dy - 2 \int_{I^2} f(x)g(y)\varepsilon_2(y) dx dy$$

$$I_a = \int_I f^2(x)\varepsilon_1(x) dx \int_I dy + \int_I \varepsilon_1(x) dx \int_I g^2(y) dy - 2 \int_I f(x)\varepsilon_1(x) dx \int_I g(y) dy =$$

$$\int_I f^2(x)\varepsilon_1(x) dx + \int_I \varepsilon_1(x) dx \geq \int_I \varepsilon_1(x) dx$$

Similarly,  $I_b \geq \int_I \varepsilon_2(y) dy$ . Thus,

$$- \int_{I^2} (f(x) - g(y))^2 C(dx, dy) \leq - \int_I \varepsilon_1(x) dx - \int_I \varepsilon_2(y) dy$$

$$\text{This implies } 2 \int_{I^2} f(x)g(y)C(dx, dy) \leq 2 - \left( \int_I \varepsilon_1(x) dx + \int_I \varepsilon_2(y) dy \right)$$

Or simply

$$\text{corr}(f, g) \leq 1 - \frac{1}{2} \left( \int_I \varepsilon_1(x) dx + \int_I \varepsilon_2(y) dy \right)$$

Because this holds for all such  $f$  and  $g$ , it holds for  $f$  and  $-g$ . Thus

$$|\text{corr}(f, g)| \leq 1 - \frac{1}{2} \left( \int_I \varepsilon_1(x) dx + \int_I \varepsilon_2(y) dy \right)$$

Provided that one of the two integrals of the right hand side is non-zero (say  $\varepsilon$ ), we obtain

$$|\text{corr}(f, g)| \leq 1 - \frac{1}{2}\varepsilon < 1.$$

The last equation shows that  $\rho_1 < 1$ , as  $\varepsilon$  is independent of  $f$  and  $g$ . Combining this with the previous theorem, we obtain exponential  $\rho$ -mixing.

### 3.3 Proof of theorem 7

Let  $X = \varphi(U)$ ,  $Y = \varphi(V)$ . Using conditional probabilities or the transformation formula, we find the joint density of  $(X, Y)$ :

$$h(x, y) = - \frac{\varphi'' \circ \varphi^{-1}(x + y)}{\left( \varphi' \circ \varphi^{-1}(x + y) \right)^3} \mathbb{I}\{x + y \leq 1\}.$$

We have

$$\begin{aligned}\rho_1(U, V) &= \sup_{f, g} \text{corr}(f(U), g(V)) = \sup_{f, g} \text{corr}(f \circ \varphi^{-1}(\varphi(U)), g \circ \varphi^{-1}(\varphi(V))) = \\ &= \sup_{\tilde{f}, \tilde{g}} \text{corr}(\tilde{f}(X), \tilde{g}(Y)) = \rho_1(X, Y).\end{aligned}$$

The second equality seems obvious, but needs clarifications. In fact the function  $\varphi^{-1}$  makes sense only for  $0 \leq x, y \leq 1$ . This condition is satisfied, because we use the density on the  $0 \leq x + y \leq 1$ . Otherwise, we would not be able to have  $\varphi^{-1} \circ \varphi(X) = X$ . The third equality holds for  $\tilde{f} = f \circ \varphi$ ,  $\tilde{g} = g \circ \varphi$ . But it is not clear that any square integrable function of  $X, Y$  can be written in this form. This equality holds, because we can write the same chain of equalities starting from  $X, Y$ , and obtain two inequalities leading to the given conclusion. Using this fact, we can therefore compute the maximal correlation coefficient of  $X, Y$ , and bound it to get a condition for exponential  $\rho$ -mixing of the Markov chain generated by the initial copula. Let  $f$  and  $g$  be such that  $\mathbb{E}(f(X)) = \mathbb{E}(g(X)) = 0$ ,  $\text{Var}(f(X)) = \text{Var}(g(Y)) = 1$ . Given the formula of the density, the correlation coefficient between  $f(X)$  and  $g(Y)$  can be computed by;

$$\text{corr}(f(X), g(Y)) = \int_0^1 \int_0^{1-x} \frac{\varphi'' \circ \varphi^{-1}(x+y) f(x) g(y)}{\left(-\varphi' \circ \varphi^{-1}(x+y)\right)^3} dy dx.$$

$$\text{Therefore, } |\text{corr}(f(X), g(Y))| \leq \int_0^1 |f(x)| \int_0^{1-x} \frac{\varphi'' \circ \varphi^{-1}(x+y) |g(y)|}{\left(-\varphi' \circ \varphi^{-1}(x+y)\right)^3} dy dx$$

$\varphi^{-1}$  is decreasing as  $\varphi$  is decreasing, and  $\varphi'$  is increasing. Thus,

$$\begin{aligned}0 &= \varphi^{-1}(1) \leq \varphi^{-1}(x+y) \leq \varphi^{-1}(x), \varphi^{-1}(y) \leq \varphi^{-1}(0) = 1; \\ \varphi'(0) &= \varphi' \circ \varphi^{-1}(1) \leq \varphi' \circ \varphi^{-1}(x+y) \leq \varphi' \circ \varphi^{-1}(x), \varphi' \circ \varphi^{-1}(y) \leq \varphi' \circ \varphi^{-1}(0) = \varphi'(1) \leq 0; \\ 0 &\leq \frac{1}{\left(-\varphi'(0)\right)^3} \leq \frac{1}{\left(-\varphi' \circ \varphi^{-1}(x+y)\right)^3} \leq \frac{1}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^{5/2} \left(-\varphi' \circ \varphi^{-1}(y)\right)^{1/2}} \leq \frac{1}{\left(-\varphi'(1)\right)^3};\end{aligned}$$

Knowing that  $U, V$  are uniform on  $[0, 1]$ , we obtain that  $X, Y$  share the same distribution with CDF and pdf

$$P(X \leq x) = P(Y \leq x) = 1 - \varphi^{-1}(x), \quad f_X(x) = f_Y(x) = \frac{1}{-\varphi' \circ \varphi^{-1}(x)}.$$

Therefore, using this density and the bound

$$h(x) = \max_{0 \leq y \leq 1-x} \varphi'' \circ \varphi^{-1}(x+y),$$

we obtain the two inequalities below, that we investigate using twice the Holder inequality and the value 1 for the variance of the random variables  $X$  and  $Y$ .

$$|\text{corr}(f(U), g(V))| \leq \int_0^1 \frac{h(x) |f(x)|}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^{5/2}} \int_0^{1-x} \frac{|g(y)| dy}{\left(-\varphi' \circ \varphi^{-1}(y)\right)^{1/2}} dx \quad (12)$$

$$|\text{corr}(f(U), g(V))| \leq \frac{1}{(\varphi'(1))^2} \int_0^1 \frac{h(x) |f(x)|}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^{1/2}} \int_0^{1-x} \frac{|g(y)| dy}{\left(-\varphi' \circ \varphi^{-1}(y)\right)^{1/2}} dx \quad (13)$$

The inequalities 12 and 13 above lead to

$$|corr(f(U), g(V))| \leq \int_0^1 \frac{h(x)|f(x)|(1-x)^{1/2}}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^{5/2}} dx \leq \left( \int_0^1 \left( \frac{h(x)}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^2} \right)^2 (1-x) dx \right)^{1/2}$$

$$|corr(f(U), g(V))| \leq \frac{1}{(\varphi'(1))^2} \int_0^1 \frac{h(x)|f(x)|(1-x)^{1/2}}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^{1/2}} dx \leq \frac{1}{(\varphi'(1))^2} \left( \int_0^1 h^2(x)(1-x) dx \right)^{1/2}.$$

Therefore, taking the supremum over all possible functions  $f$  and  $g$ , we obtain in both cases:

$$\rho_1^2 \leq \int_0^1 \left( \frac{h(x)}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^2} \right)^2 (1-x) dx \quad \text{or} \quad \rho_1^2 \leq \frac{1}{(\varphi'(1))^4} \int_0^1 h^2(x)(1-x) dx.$$

So,  $\rho_1 < 1$  if

$$\int_0^1 \left( \frac{h(x)}{\left(-\varphi' \circ \varphi^{-1}(x)\right)^2} \right)^2 (1-x) dx < 1 \quad \text{or} \quad \int_0^1 h^2(x)(1-x) dx < (\varphi'(1))^4 \quad \text{for} \quad \varphi'(1) \neq 0.$$

Now, if  $\varphi''$  is decreasing, then  $h(x) = \varphi'' \circ \varphi^{-1}(1) = \varphi''(0)$ , and if  $\varphi''$  is increasing, then  $h(x) = \varphi'' \circ \varphi^{-1}(x)$ . This takes into account the fact that  $\varphi^{-1}$  is decreasing, and ends the proof of the theorem.

In general, the function  $h(x)$  is easy to compute because it needs only the behavior of the second derivative of the generator on the interval of definition. The use of the standard generation helps to avoid question about the ranges and domains of function that are used throughout the proof.

### 3.4 Proof of corollaries 9, and 10

For corollary 9, the existence of a spectral gap in  $L^2(\chi, \mu)$  is equivalent to  $L^2(\chi, \mu)$ -geometric ergodicity for reversible Markov chains. So, the Markov chain generated by  $P$  is  $L^2(\chi, \mu)$ -geometrically ergodic. Therefore, a straight application of lemma 8 leads to the conclusion.

For corollary 10, according to [11, theorem 2.1], both Markov processes have a spectral gap in  $L^2(\chi, \pi)$ . The lemma states, that therefore, their convex combination has a spectral gap in  $L^2(\chi, \pi)$ . Applying again [11, theorem 2.1], the spectral gap is equivalent to geometric ergodicity; so, the Markov chain generated by the convex combination is geometrically ergodic.

### 3.5 Proof of result 2.4.1

**Proof** The proof of the formula of the  $n$  steps transition copula is based on the multiplicative property of the copula family and the recurrence relationship that can be easily established between  $n$  and  $n+1$  steps copulas. Moreover, any set  $S_{u_0} = [0, u_0]$ ,  $u_0 \leq 1/2$ , is a small set for  $\alpha + \beta \neq 1$ , as for any  $u_1, u_2 \leq u_0$ ,  $S_{u_1} \subset S_{u_0}$  and  $S_{u_2} \subset S_{u_0}$ ,

$$P(X_n \in S_{u_1}, X_0 \in S_{u_2}) = C^n(u_1, u_2) \geq (1 - \alpha_n - \beta_n)u_1u_2 = (1 - \alpha - \beta)P(X_n \in S_{u_1})P(X_0 \in S_{u_2}).$$

Now,

$$\begin{aligned} |P(X_n \in S_{u_0} | X_0 \in S_{u_0}) - P(X_0 \in S_{u_0})| &= \left| \frac{C^n(x_0, x_0) - x_0^2}{x_0} \right| \\ &= \frac{\alpha_n M(u_0, u_0) - (\alpha_n + \beta_n)u_0^2 + \beta_n W(u_0, u_0)}{u_0} \\ &= |\alpha_n - (\alpha_n + \beta_n)u_0|; \text{ take } u_0 = 1/2; \end{aligned}$$

$$|P(X_n \in S_{u_0} | X_0 \in S_{u_0}) - P(X_0 \in S_{u_0})| = \frac{\alpha_n - \beta_n}{2} = \frac{1}{2}|\alpha - \beta|^n.$$

So, we conclude that this goes to zero exponentially fast for  $-1 < \alpha - \beta < 1$ ; for  $\alpha = 1$  (it corresponds to the copula  $M$ ) or  $\beta = 1$  (it corresponds to the copula  $W$ ), the Markov process is not a  $\beta$ -mixing. For any convex combination of the two copulas (corresponding to  $\alpha + \beta = 1$ ), we can't use this method to provide the answer, as the coefficient in the inequality defining the small set is zero, and thus, the sets we are using are not small sets. On the other hand, the problem can still be handled. The corresponding transition operator acts as follows:

$$Qf(u) = \alpha f(u) + (1 - \alpha)f(1 - u).$$

Therefor, if we can find a function  $f$  defined on  $I$ , such that  $\mathbb{E}(f) = 0$  and  $f(1 - u) = f(u)$  for all  $u$ , then will have  $Qf(u) = f(u)$ , leading to  $\|Qf\|_2 = \|f\|_2$  and  $\|Q\| = 1$ . There is a huge class of such functions. For instance,  $f_n(x) = \cos(2n\pi x)$  satisfies these assumptions. Thus,  $\|Q\| = 1$  - there is no spectral gap. Therefore, we can conclude that we have exponential convergence for all parameters except when  $\alpha + \beta = 1$ . By [11, theorem 2.1], the Markov process generated by a Frechet copula (or Mardia copula) is therefore an exponential  $\rho$ -mixing and exponential  $\beta$ -mixing, except when they represent a convex combination of  $M$  and  $W$ .

### 3.6 Proof of theorem 2

We consider the functional operator  $T : L_2([0, 1]) \rightarrow L_2([0, 1])$  defined by:

$$T(f)(x) = \int_0^1 f(y)c(x, y)dy, \quad c(x, y) = C_{12}(x, y).$$

Define the operator  $T_1 f(x) = T(f(x) - E(f(x)))$ . The following holds:

$$\|T\| = \sup_{f \in L_2^0([0, 1])} \frac{\|T(f(x))\|_2}{\|f(x)\|_2} = \sup_{f \in L_2([0, 1])} \frac{\|T_1(f(x))\|_2}{\|f(x) - E(f(x))\|_2}. \quad (14)$$

Let  $H = L_2([0, 1])$  be a Hilbert space. Let  $T$  be a bounded operator defined on  $H_0 = L_2^0([0, 1])$ , and  $\{e_i(x), i \in \mathbb{N}\}$  be an orthonormal basis of  $H$  such that  $E(e_i(x)) = 0$  for all  $j \neq 1$ . Then

$$\|T\|^2 \leq \sum_{i>1} \|T(e_i)\|_2^2.$$

**Proof:**

$$E(f(x)) = 0 \text{ and } E(e_i(x)) = 0 \text{ for } i > 1 \longrightarrow a_1 = 0.$$

$$f(x) = \sum a_i e_i(x) \longrightarrow \|f\|_2 = (\sum_{i>1} a_i^2)^{1/2};$$

$$\|Tf\|_2 = \|(\sum_{i>1} a_i T e_i)\|_2 \leq (\sum_{i>1} a_i^2)^{1/2} (\sum_{i>1} \|T e_i\|_2^2)^{1/2};$$

The last inequality above uses the Holder or Cauchy inequality. This leads to  $\|T\| \leq (\sum_{i>1} \|T e_i\|_2^2)^{1/2}$ . We will bound this norm for the operator that we have defined to prove the theorem, and find a sufficient condition for the norm of the initial operator to be less than 1. One of the most convenient basis is the following:

$$1, e_i = \sqrt{2} \sin(2i\pi x); e_i = \sqrt{2} \cos(2i\pi x); i = 1, 2, \dots$$



For this basis, we have  $E(e_i(x)) = 0$  except for  $e_i = 1$  and it's an orthonormal basis of our Hilbert space. It remains to evaluate or bound the quantities  $\|Te_i\|_2^2$ .

**Case 1.**  $e_i = \sqrt{2}\cos(2i\pi x)$ .

$$(1/\sqrt{2})Te_i(x) = \int_0^1 c(x, y)\cos(2\pi iy)dy = c(x, y)\frac{\sin(2i\pi y)}{2i\pi}\Big|_{y=0}^1 - \frac{1}{2i\pi} \int_0^1 c_y(x, y)\sin(2\pi iy)dy$$

Therefore, for the absolute value of  $(1/\sqrt{2})Te_i(x)$ , using  $|\sin(2\pi iy)| \leq 1$  and (2), we obtain:

$$|(1/\sqrt{2})Te_i(x)| \leq \frac{1}{2i\pi} \int_0^1 |c_y(x, y)|dy$$

$$\|Te_i\|_2^2 \leq \frac{1}{2(i\pi)^2} \left\| \int_0^1 |c_y(x, y)|dy \right\|_2^2;$$

$$\text{So } \|Te_i\|_2^2 \leq \frac{k_1}{2(i\pi)^2}.$$

**Case 2.**  $e_i = \sqrt{2}\sin(2i\pi x)$ . Same as above with the only difference that for this case, the first part is not zero, but  $-(\frac{1}{2i\pi})[c(x, 1) - c(x, 0)]$ . So,

$$|(1/\sqrt{2})Te_i(x)| \leq \frac{1}{2i\pi} [|c(x, 1) - c(x, 0)| + \int_0^1 |c_y(x, y)|dy].$$

$$\|Te_i\|_2^2 \leq \left\| \frac{1}{2i\pi} [|c(x, 1) - c(x, 0)| + \int_0^1 |c_y(x, y)|dy] \right\|_2^2.$$

$$\|Te_i\|_2^2 \leq \frac{k_2}{2(i\pi)^2}.$$

Taking into account both cases, using

$$\sum_{i>0} \frac{1}{i^2} = \frac{\pi^2}{6},$$

$$\sum_{i>1} \|T(e_i)\|_2^2 = \frac{k_1 + k_2}{2(\pi)^2} \sum_{i>0} \frac{1}{i^2} = \frac{k_1 + k_2}{12}.$$

This norm is less than 1 for  $k_1 + k_2 < 12$ , and therefore the generated Markov chain is an exponential  $\rho$ -mixing.

## 4 Acknowledgments

The author thanks his advisor Dr Magda Peligrad for her support in life, for her questions and for her help in the preparation of this article.

## References

- [1] Beare, B.K. (2010). Copulas and Temporal Dependence. *Econometrica*, Volume 78, Issue 1, pages 395–410.
- [2] Bradley, R.C.(2007). Introduction to strong mixing conditions. V 1. Kendrick press.

- [3] Clayton, D. G.(1978). A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incident. *Biometrika*, 65, 141-151.
- [4] Darsow, W. F. , Bao Nguyen and Elwood T. Olsen .(1992) Copulas and Markov processes. *Illinois journal of mathematics*. v. 36,No.4, Winter 1992.
- [5] Doukhan, P., Massart, P., Rio, E. (1994) The functional central limit theorem for strongly mixing processes. *Annales de l'I.H.P., section B*, tome 30, №1, pp.63-82.
- [6] Embrechts,P., Lindskog,F., McNeil ,A. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, ed. S. Rachev, Elsevier, 329-384.
- [7] Longla M., Peligrad M (2011). Some Aspects of Modeling Dependence in Copula-based Markov chains. *Journal of Multivariate Analysis*, <http://dx.doi.org/10.1016/j.jmva.2012.01.025>
- [8] Nelsen R.B.(2006). An introduction to copulas. 2nd ed. Springer, New York.
- [9] Peligrad M. (1997) On the central limit theorem for triangular arrays of  $\phi$ -mixing sequences. *Asymptotic methods in statistics and probability* (Ottawa, Ontario ) pp.49-55.
- [10] Peligrad M. (1993) Asymptotic results for  $\phi$ -mixing sequences. *Doebelin and Modern probability* (H. Cohn ed.) *Contemporary Mathematics* 149 pp. 171-177. American Mathematical Society; Providence, Rhodes Island.
- [11] Roberts G.O., Jeffrey S.R. (1997). Geometric ergodicity and hybrid Markov chains. *Electronic communications in probability*. 2, 13-25.
- [12] Chen X., Wei Biao Wu, Yanping Yi.(2009). Efficient estimation of copula-based semi-parametric Markov models. *The annals of statistics*. Vol. 37. No. 6B, 4214-4253.